

## 1.0 EVALUATION OF LABORATORY REPLICATES AND FIELD DUPLICATES

---

For individual chemicals and sums, the process explained in the attached flow chart, Figure 1-1, which is generally consistent with EPA comments and method agreements, was used to further evaluate duplicate outliers in stormwater. Table 1-1 lists all duplicates with Relative Percent Difference (RPD) values exceeding the levels presented in Table 4.2 of the Portland Harbor RI/FS Round 2 QAPP Round 3A Stormwater Sampling (Integral 2007). Since no RPD limit was specified for PCBs, the screening level RPD for phthalates, pesticides, and PAHs was used, which is 30 percent.

A detailed evaluation regarding the treatment of replicate/duplicates is presented in Table 1-1, along with the rationale for the recommended replicate/duplicate handling following the decision process shown in Figure 1-1. For this preliminary screening process, all non-detect results were included at one half the detection limit.

The screening resulted in 80 parent and replicate/duplicate pairings out of approximately 500 total pairings having an RPD greater than the screening factor and therefore retained for further evaluation as presented in Figure 1-1. Through the additional analysis, 24 parent and replicate/duplicate pairings were subjected to an additional segregation evaluation, which is approximately 5 percent. In other words, in 95 percent of the cases, replicates or duplicates were averaged per standard RI database rules.

Additionally, out of the 24 pairs summarized in the attached Table 1-1, eight of these pairs are from OF-18, which is a multiple land use site. Data from multiple land use sites were collected with the intent to perform an uncertainty analysis and are not used directly in any loading calculations. Therefore, these samples are not further discussed here, but are included in Table 1-1 for reference and are discussed further in Section E3.7.2.

Out of the remaining 16 cases, only one pair was completely segregated (removed) from the stormwater loading working database. In the other 15 cases, either the parent or the replicate/duplicate was segregated and the other half of the pair was retained in the working database.

Due to the limited data set for sediment traps, all sediment trap duplicates were averaged with parent samples.

## 2.0 RECLASSIFICATION ANALYSIS

The *a-priori* non-representative locations and chemicals are shown below in Table 2-1.

Table 2-1. Chemicals and Sites for Further Analysis

Outfall #	Facility/Location	Non-Representative Chemicals for Further Analysis
WR-22	OSM	PCBs, PAHs, metals
WR-123	Schnitzer International Slip	PCBs, phthalates, metals
WR-384	Schnitzer - Riverside	Metals, PCBs
WR-107	GASCO	PAHs
WR-96	Arkema	Pesticides
WR-14	Chevron - Transportation	PAHs
WR-161	Portland Shipyard	PAHs, phthalates, metals, PCBs
WR-4	Sulzer Pump	PAHs, metals, PCBs
WR-145	Gunderson	PCBs, PAHs, phthalates, metals
WR-147/148	Gunderson (former Schnitzer)	Phthalates, metals, PCBs, PAHs
Within OF-17	GE	PCBs
WR-183/Basin R	Terminal 4 - Slip 1	PAHs, TOC
WR-181/Basin Q	Terminal 4 - Slip 1	Metals, PAHs, TOC
WR-177/Basin M	Terminal 4 - Slip 1	Metals, PAHs
WR-169/BasinD	Terminal 4	Metals, PAHs
WR-20/Basin L	Terminal 4 - Wheeler Bay	PAHs
OF-22B	City -Doane Lake Industrial Area	Pesticides, Metals
WR-510	St. Johns Bridge/Highway 30	PCBs, others (bridge repaving activity)

For individual chemicals and sums, the process explained in the attached flow chart, Figure 2-1, was used to evaluate the classification of data. Sites with both Heavy Industrial and Light Industrial land use types were evaluated. There were no sites in the Residential and Open Space land uses identified for further analysis; therefore, the reclassification analysis does not address these land uses. St. Johns Bridge data will be examined as part of a separate process and is not included in the RI data set.

As shown on Figure 2-1, the evaluation process first assessed whether representative locations should remain representative or become non- representative, and then assessed whether non-representative sites were better categorized as representative. The results are summarized in Tables 2-2a–f, which provides the decisions made for Step 1 and Step 2 and the resulting recommended final categorization.

Each chemical within each chemical group was evaluated separately to determine final categorizations, with the exception of PCBs. PCB sites were classified as representative or non-representative on the basis of the entire set of congeners and total PCBs, and therefore a site could not be non-representative for one congener and representative for another congener.

### 3.0 CALCULATION OF LOADS

---

A full range of summary statistics, including central tendencies and confidence limits were generated such that evaluations of various types of loading estimate scenarios and modeling sensitivity analyses can be supported. Loads were estimated as a product of the calculated concentration estimates and the flow rate from the 50th percentile flow year. The annual mass loads were generated by adding the loading contributions from each land use and non-representative site for each fate and transport model segment:

Annual stormwater chemical load (kg/yr) = heavy industrial stormwater chemical load (kg/yr) + light industrial stormwater chemical load (kg/yr) + residential stormwater chemical load (kg/yr) + parks/open space stormwater chemical load (kg/yr) + major transportation stormwater chemical load (kg/yr) + “non-representative” site stormwater chemical load (kg/yr)

To express the uncertainty in these estimates, a range of loading rates was calculated for each indicator contaminant, including the geometric mean, basin-weighted mean, 5<sup>th</sup> percentile, and 95th percentile, in order to show a representative range of concentrations.

### 3.1 COMPOSITE WATER

---

#### 3.1.1 Minimum Sample Size Requirements

Summary statistics generated were often based on data sets with few observations and/or detected values. Hypothesis testing (goodness-of-fit), interpolation (i.e., Regression on Order Statistics [ROS]), and estimation (upper confidence limits [UCL]) methods used to generate summary statistics may not be appropriate or reliable due to the uncertainty in the representativeness of the data set for the population of interest. In addition, ProUCL has incorporated minimum sample size requirements into the statistical routines and may not provide such statistics or test results for small data sets. As such, the following decision rules, based on both statistical principles and recommendations provided in the ProUCL Version 4.0 Technical Guide and User Guide (USEPA 2007) and practical limits of the ProUCL software, were used to determine whether specific test results or statistics would be calculated and presented in summary statistics. The decision rules are as follows:

- For analyte/matrix/land-use combinations with  $5 \leq N < 8$ , advanced summary statistics were generated and presented but should be interpreted with caution due to the limited number of samples.
- For analyte/matrix/land-use combinations with  $5 \leq N < 10$ , bootstrap methods for estimating UCL were avoided due to uncertainties in the bootstrapping operation with low sample numbers; ProUCL recommends a minimum of 10 to 15 samples for bootstrapping operations.

- For analyte/matrix/land-use combinations with less than four detected observations, goodness-of-fit (GOF), ROS, and bootstrap operations are unreliable and were not used. ProUCL will not generate GOF, ROS-based summary statistics, Kaplan-Meier and bootstrap estimates for this scenario.

### 3.1.2 Summarize Unweighted Composite Water Data

Summary statistics on unweighted data were calculated by land use and are presented in Table E3-2.

#### 3.1.2.1 Summary Statistics for Data Averaged by Site, and Then Pooled by Chemical and Land Use

1. The lognormal ROS method was used to impute non-detect values using the pooled data set. Estimation of non-detects was necessary in this step in order to estimate the averages by site, because sample numbers or detected sample numbers were too small on a per-site basis to use other techniques (i.e., ROS or Kaplan-Meier) to interpolate non-detects or estimate averages for each basin. A lognormal distribution was used in the ROS estimates. In cases where the ROS method was unreliable due to a limited number of samples or limited detected samples as described in Section 3.1, half the detection limit was substituted for non-detect values.
2. The data were averaged by site, in order to come up with one value for each sample location.
  - a. ProUCL was used to generate summary statistics on the averaged by site data consistent with recommendations for such statistics provided in the ProUCL Version 4.0 Technical Guide and User Guide (USEPA 2007). Statistics of interest for the RI Report include geomean, and the 5<sup>th</sup> and 95<sup>th</sup> percentiles.

### 3.1.3 Summarize Weighted Composite Water Data

Summary statistics on data averaged by site and weighted using a unit flow factor are also presented in Table E3-2 in the column labeled “Basin Weighted Mean Concentration.” The steps for this calculation were as follows:

1. Using the same data set created above with substituted values and data averaged by site, the data were weighted using the following method:

$C_{\text{weighted}} = C \times W \times N$ , where

$C$  = the average concentration from each sample location

$W$  = weighting factor, a unitless factor for each sample location based on its runoff volume divided by the sum of all volumes for all locations

$N$  = the number of sample locations in a land use category

2. Microsoft Access was used to calculate the mean concentration for each land use.

### 3.1.4 Summary Statistics for Non-Representative Sites

1. Since it was not possible to use ROS or Kaplan-Meier to calculate means for each individual sample location due to limited samples and non-detects, half the detection limit was substituted for non-detects.
2. The mean value for each non-representative site is presented in Table E3-2. Note that since only one value exists for each site after averaging the data by site, only one value is presented.

## 3.2 SEDIMENT TRAP DATA

---

Sediment trap data were collected during both Rounds 3A and 3B Stormwater Sampling. The purpose of Round 3B sampling was to fill data gaps where data were not collected in the first round. However, there are a few instances where the same analyte was measured at the same location in both Round 3A and 3B. This occurred if a limited sample mass collected during Round 3A led to elevated detection limits. Thus, it could be expected that some non-detect values occurred in Round 3A due to limited sample mass. In this case, the analytes were measured again in Round 3B if sufficient sediment was available.

There were 16 instances where there was a non-detect sample collected for a particular analyte during both Round 3A and Round 3B. In most cases, the non-detect value in Round 3A and 3B were similar. However, in the case of three pesticide samples collected at OF-49, the non-detect values in Round 3A were 10 times greater than the non-detect samples collected in Round 3B. In the case of these three samples, the high non-detect samples collected in Round 3A were segregated.

In all other cases, samples for sediment traps were averaged prior to calculation of any statistics. It should be noted that this procedure differs from the treatment of the composite water samples. For sediment traps, if two detected samples exist for a particular sampling location, then samples were averaged. If there was one non-detect and one detect sample, then the detect sample was retained, and the non-detect sample was segregated. If both samples were non-detect, then the samples were averaged and the non-detect qualifier remained except in the three cases discussed above.

### 3.2.1 Step 1 – Summary Statistics for Sediment Trap Data for Representative Land Use Sampling Locations

The process for calculating unweighted statistics on the data is explained below.

1. After averaging of samples by location, data for each land use were reformatted to meet ProUCL requirements. Records identified as non-representative were treated as independent data sets on a chemical- and location-specific basis and are discussed below. Statistics were only calculated to the extent practical in accordance with ProUCL guidance due to the small data set associated with each.
2. ProUCL was used to conduct graphical and statistical (i.e., GOF) tests to determine the underlying data distribution (or lack thereof) for each analyte and

land use.

3. ProUCL was used to generate summary statistics for each land use consistent with recommendations for such statistics provided in the ProUCL Version 4.0 Technical Guide and User Guide (USEPA 2007). Statistics of interest include the geomean and the 5<sup>th</sup> and 95<sup>th</sup> percentiles.
4. After calculating statistics, the chemical solids loading rate (a concentration in terms of mass per volume water) similar to that obtained via stormwater was calculated and is presented in Table E3-2. The measured sediment concentration statistics (Cs in µg/kg) were multiplied by a central tendency (i.e., geometric mean) of the total suspended solids concentration (TSS in kg/L) measured in composite water for a particular land use to get a concentration in terms of (µg/L)

### **3.2.2 Step 2 – Summary Statistics for Weighted Sediment Trap Data for Representative Land Use Sampling Locations**

The process for calculating weighted statistics on the data is as follows:

1. Using the sediment trap data set with samples averaged by site as discussed above, the lognormal ROS method was used to impute non-detect values. A lognormal distribution was used in the ROS estimates for the following reasons: 1) the normal ROS estimation method frequently imputes negative values for non-detects, which is not possible, and 2) environmental data frequently assumes a lognormal distribution; there is an underlying assumption of lognormality for these stormwater data. In cases where the ROS method was unreliable due to limited samples or limited detected samples as described in Section 3.1, half the detection limit was substituted for non-detects.
2. The data were weighted using the following method:

$$C_{\text{weighted}} = C \times W \times N, \text{ where:}$$

C = the average concentration from each sample location

W = weighting factor, a unitless factor for each sample location based on its runoff volume divided by the sum of all volumes for all locations

N = the number of sample locations in a land use category

3. ProUCL was used to generate summary statistics on the weighted data consistent with recommendations for such statistics provided in the ProUCL Version 4.0 Technical Guide and User Guide (USEPA 2007). Statistics of interest include the geometric mean, 5<sup>th</sup> and 95<sup>th</sup> percentiles.
4. After calculating statistics, the chemical solids loading rate was calculated exactly as described above for unweighted data and are presented in Table E3-2.

### **3.2.3 Step 3 – Sediment Trap Data for Non-Representative Land Use Sampling Locations**

1. There is generally only one data point for each non-representative sampling

location, so statistical methods cannot be used to calculate substitution values for non-detects. Therefore, half the detection limit was substituted for non-detects.

2. There is generally only one data point for each sampling location and chemical so no statistics were calculated.
3. The chemical solids loading rate (concentration in water terms) was then calculated from the single value available at each site in the same manner as noted for representative data above and these values are presented in Table E3-2.